



TITLE:

大域結合写像としてのニューラルネットワークモデルとその応用(研究会「複雑系」,研究会報告)

AUTHOR(S):

野沢, 浩

CITATION:

野沢, 浩. 大域結合写像としてのニューラルネットワークモデルとその応用(研究会「複雑系」,研究会報告). 物性研究 1992, 59(3): 340-341

ISSUE DATE:

1992-12-20

URL:

<http://hdl.handle.net/2433/95004>

RIGHT:

大域結合写像としてのニューラルネットワークモデルとその応用

(株) ニコンシステム 野沢 浩

1 大域結合写像としてのニューラルネットワークモデル

Hopfieldによって提案されているニューラルネットワークモデル[1]は、ニューロン i ($i = 1, \dots, M$)の入力に対する連続時間の力学系として次のように定義される。

$$\frac{du_i(t)}{dt} = \sum_{j=1}^M T_{ij}v_j(t) - \frac{u_i(t)}{R} + I_i \quad (1)$$

$$v_i(t) = f(u_i(t)) = \frac{1}{2} \left(1 + \tanh\left(\frac{u_i(t)}{2a}\right) \right) \quad (2)$$

ここで、 $u_i(t)$ はニューロン i の連続時間 t における入力を表し、 $v_i(t)$ ($0 < v_i(t) < 1$)はニューロン i の連続時間 t における出力を表し、 T_{ij} はニューロン j からニューロン i へのシナプス荷重を表し、 I_i はニューロン i のしきい値を表し、 R ($R > 0$)はニューロンの減衰特性を表し、 a ($a > 0$)はニューロンのしきい特性を表す。

負の自己結合 $T_{ii} = -T$ ($T > 0$)を持つモデル(1)-(2)をオイラー法により差分化することで、次のニューロン i の入力に対する離散時間の力学系を得る。

$$u_i(n+1) = \left(1 - \frac{\Delta t}{R}\right)u_i(n) - \Delta t T f(u_i(n)) + \Delta t \sum_{j=1}^M T_{ij}f(u_j(n)) + \Delta t I_i \quad (3)$$

ここで、 $u_i(n)$ はニューロン i の離散時間 n における入力を表し、 Δt は差分ステップを表す。また、モデル(3)と等価な系として次のニューロン i の出力に対する離散時間の力学系を考えることもできる。

$$v_i(n+1) = f\left(-\Delta t T \sum_{k=0}^n \left(1 - \frac{\Delta t}{R}\right)^k v_i(n-k) + \Delta t \sum_{j=1}^M T_{ij} \sum_{k=0}^n \left(1 - \frac{\Delta t}{R}\right)^k v_j(n-k) + R I_i\right) \quad (4)$$

ここで、 $v_i(n)$ ($0 < v_i(n) < 1$)はニューロン i の離散時間 n における出力を表す。また、ニューロンの入力と出力の間には次の関係が成り立つ。

$$v_i(n) = f(u_i(n)) \quad (5)$$

これらのモデル(3)、(4)は合原等によって提案されているカオスニューラルネットワークモデル[2]の中で、最も単純なものを与える。

モデル(4)を金子によって精力的に調べられている大域結合写像(Globally Coupled Map; GCM)[3]として表現する為に、次の新しい変数ニューロン i の内部保持値(仮想変数) $p_i(n)$ ($0 < p_i(n) < 1$)を導入する。

$$p_i(n) = \frac{\Delta t}{R} \left(\sum_{k=0}^{n-1} \left(1 - \frac{\Delta t}{R}\right)^k v_i(n-1-k) \right) \quad (6)$$

モデル(4)へ式(6)を適用することにより、ニューロン i の内部保持値に対する力学系は次式のような大域結合写像として定義される。

$$p_i(n+1) = F_{q_i(n)}(p_i(n)) \quad (7)$$

$$q_i(n) = \frac{1}{T} \left(\sum_{j=1}^M T_{ij} p_j(n) + I_i \right) \quad (8)$$

$$F_q(p) = \left(1 - \frac{\Delta t}{R}\right)p + \frac{\Delta t}{R} \left\{ 1 - f(RT(p - q)) \right\} \quad (9)$$

ここで、 $q_i(n)$ はニューロン j ($j \neq i$)からニューロン i への寄与を表す変数であり、この変数により内部保持値 $p_i(n)$ の従う局所的な場(非線形写像)が決定される。また、ニューロン i の出力は次式のように計算される。

$$v_i(n) = 1 - f\{RT(p_i(n) - q_i(n))\} \quad (10)$$

式(8)においてシナプス荷重 T_{ij} が $T_{ij} = -T\delta_{ij}$ と与えられると仮定することで、単一ニューロン i の内部保持値に対する次の単純な1次元写像を得ることができる。

$$p_i(n+1) = F_{q_i}(p_i(n)) \quad (11)$$

$$q_i = \frac{I_i}{T} \quad (12)$$

ここで、 q_i は1次元写像(11)の制御パラメータを表す。

ニューラルネットワークモデル(7)-(9)と単一ニューロンモデル(11)-(12)の内部保持値に対する力学系を比較することにより、モデル(7)-(9)の大域結合写像としての特徴は内部保持値の従う非線形写像の多様性(Variety of Maps)にあることがわかる。これにより、単一ニューロンモデルのときには内部保持値 $p_i(n)$ は時刻 n に依らないただ一つの非線形写像 F_{q_i} により状態を更新される(図1a参照)が、ニューラルネットワークモデルのときには $p_i(n)$ は時刻 n 毎に異なる非線形写像 $F_{q_i(n)}$ により状態を更新される(図1b参照)ことになる。

モデル(7)-(9)において、写像の多様性を利用することにより、動的な情報処理モデルを実現することが可能となる。我々が行った動的な情報処理モデルによる工学的応用例として次の2つ(a)曖昧キーワードによる情報検索、(b)巡回セールスマン問題(Traveling Salesman Problem; TSP)の解法がある[5]。

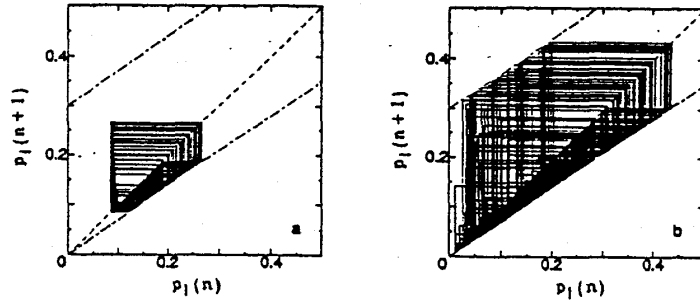


図1: 内部保持値 $p_i(n)$ の軌道の振舞い($\Delta t/R = 0.3$, $a/RT = 0.006$, $q_i = 0.09$)
(a)単一ニューロンモデル, (b)ニューラルネットワークモデル

2 動的情報処理モデルによる巡回セールスマン問題の解法

巡回セールスマン問題は「一人のセールスマンが N 個の都市を訪問して商売をすることになっている、各都市間の距離が与えられたとして、各都市を一度ずつ訪れ最後に出発点に戻ってくるような経路(訪問の順序)の中で最短となるものを求めよ」で与えられる、解法が困難(NP完全)な離散の組み合わせ最適化問題である。

HopfieldとTankは、モデル(1)-(2)によるこの問題の具体的な解法例(都市数 $N = 10$)を示した[4]。しかし、彼等のモデルは勾配系である為に、得られるTSPの解が初期条件に依存してしまう問題点があった。これに対して我々はモデル(7)-(9)をカオス状態に設定することで、動的情報処理モデルを実現しTSPの解法を行った。ニューロンを用いたTSPの解の表現方法はHopfieldとTankが示した方法を用いた。また評価関数は制約条件項と距離反映項により次式のように定めた。

$$E(n) = \frac{1}{2} \left[A \left\{ \sum_{i=1}^N \left(\sum_{k=1}^N v_{ik}(n) - 1 \right)^2 + \sum_{k=1}^N \left(\sum_{i=1}^N v_{ik}(n) - 1 \right)^2 \right\} + B \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N d_{ij} v_{ik}(n) (v_{jk+1}(n) + v_{jk-1}(n)) \right] \quad (13)$$

ここで、 A, B は正の定数を表す。評価関数 $E(n)$ によりTSPを解法するためのシナプス荷重としきい値は次のように与えられる。

$$T_{ikj} = -A(\delta_{ij}(1 - \delta_{kl}) + \delta_{kl}(1 - \delta_{ij})) - Bd_{ij}(\delta_{lk+1} + \delta_{lk-1}) \quad (14)$$

$$I_{ik} = A \quad (15)$$

ここで、 δ_{ij} はクロネッカーのデルタを表す。

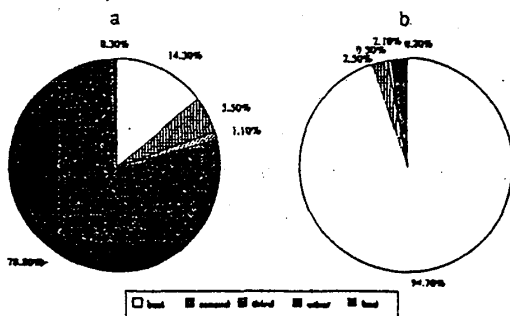


図2: TSPの解法能力

(a)モデル(1)-(2), (b)モデル(7)-(9) ($T = 1.3$)

図2にモデル(1)-(2)と(7)-(9)のTSPの解法能力を示す。解法にあたっては、都市の配置データとしてHopfield等のオリジナルデータを用い、定数 $A = B = 1$ とした。この場合、明かにモデル(7)-(9)の解法能力が優れている。

参考文献

- [1] J. Hopfield, Proc. Natl. Acad. Sci. USA 79(1982)2554.
- [2] K. Aihara, T. Takabe and M. Toyoda, Phys. Lett. A 144(1990)333.
- [3] K. Kaneko, Physica D 41(1990)137.
- [4] J. Hopfield and D. W. Tank, Biol. Cybern. 52(1985)52.
- [5] H. Nozawa, preprint.